BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 6: e94420
doi: 10.3897/biss.6.94420

OPEN ACCESS

Conference Abstract

Diversifying the GBIF Data Model

Tim Robertson‡, John R Wieczorek§, Mélianie Raymond‡

‡ Global Biodiversity Information Facility, Copenhagen, Denmark
§ Global Biodiversity Information Facility, Bariloche, Argentina

Corresponding author: Tim Robertson (trobertson@gbif.org)

Received: 02 Sep 2022 | Published: 07 Sep 2022

Citation: Robertson T, Wieczorek JR, Raymond M (2022) Diversifying the GBIF Data Model. Biodiversity Information Science and Standards 6: e94420. https://doi.org/10.3897/biss.6.94420

Abstract

The Global Biodiversity Information Facility (GBIF) has been immensely successful in mobilizing a large number of records documenting species occurrence through a global network of data publishers. Today over 2 billion records are available for search and download through the GBIF.org infrastructure, with over 200 million originating in natural science collections or gleaned from citations of specimens in literature. These data are most commonly made available through the TDWG standards *Darwin Core* and *Access to Biological Collections Data (ABCD)* with their various extensions.

The growth of the GBIF network has generated new challenges and a desire for GBIF to accommodate more varied types of data (e.g., ecological, eDNA, biological interactions, material samples with subsamples and associated media, taxonomic treatments), acknowledging that biodiversity data are more complex than 'just' the occurrence of species in time and space. GBIF has initiated work to explore the diversification of the data model that underpins the datasets connected to GBIF, and the services provided by its infrastructure. During 2021, several case studies were drafted and analyzed to establish a draft common domain model that could satisfy the key aspects of the studies. The cases cover a range of concerns, including automated monitoring, identification through sequencing techniques, biotic interactions, environmental measurements, and improvements to the handling of data representing preserved material. The draft model has been presented publicly several times and is being tested and matured to accommodate real datasets during 2022.

© Robertson T et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

This presentation will summarise the process followed, provide an overview of the cases that have been explored, and show the resulting model. We will describe the ongoing opportunities for community engagement as this work continues. We will provide insight into some of the lessons we have learned and how this research has the potential to help integrate many efforts relating to the use of the Darwin Core-related standards today.

Keywords

species, standards, biodiversity, evidence, occurrence

Presenting author

John R Wieczorek and Tim Robertson

Presented at

TDWG 2022